Solution 1: Hard Margin Classifier

(a) The dataset is linearly separable. A hard margin SVM is suitable as a solution.

- The safety margin γ should be as large as possible subject to the constraint equations:

$$y^{(i)}\left(\left\langle \boldsymbol{\theta}, \mathbf{x}^{(i)} \right\rangle + \theta_0\right) \geq 1$$

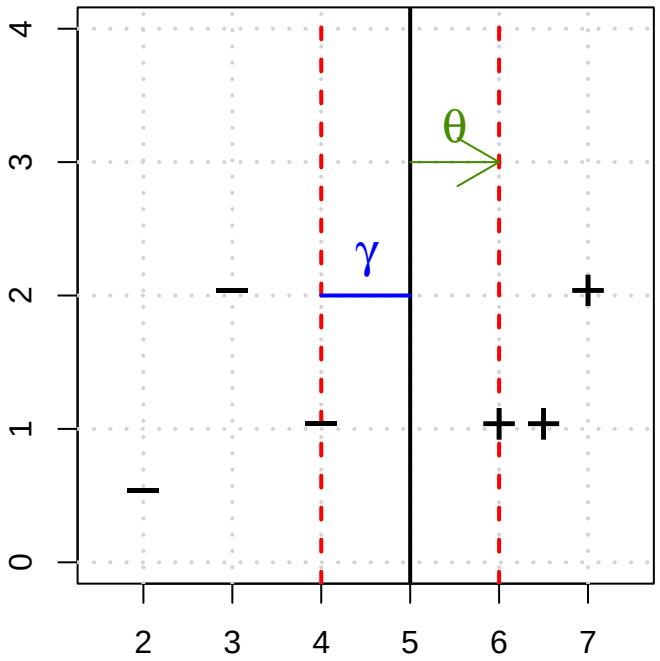
  We can derive graphically that the separating hyperplane lies between the points $(4,1)^T$ and $(6,1)^T$. The maximum margin is achieved when the hyperplane lies exactly between these two points, with a value of $\gamma = 1$.
- The norm of $\boldsymbol{\theta}$ can be calculated as :

$$\|\boldsymbol{\theta}\| = \frac{1}{\gamma} = \frac{1}{1} = 1$$

- By inspection, the separating hyperplane equation is given by:

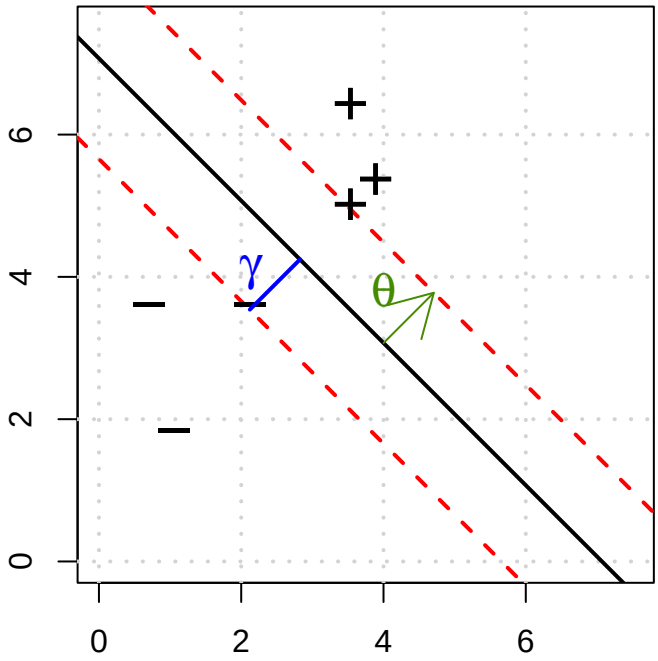
$$x_1 = 5$$

  Using the formula $\boldsymbol{\theta}^T\mathbf{x} + \theta_0 = 0$, we obtain that $\theta_1 = 1$, $\theta_2 = 0$ and $\theta_0 = -5$
- The support vectors are the ones that determine the margins, in this case: $(4,1)^T$ and $(6,1)^T$



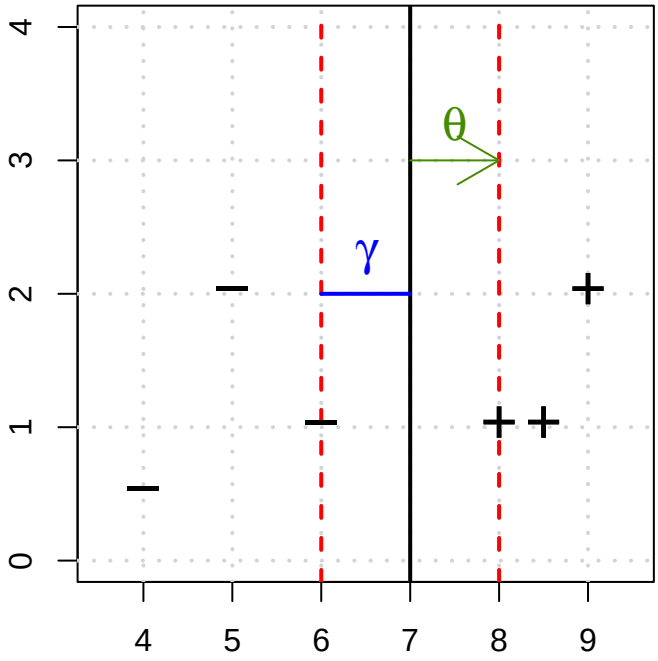
(b) Let's see what happens to the points calculated in (a) if we apply some changes to the problem.

- If the points are rotated 45 degrees counterclockwise, the problem now looks like this:



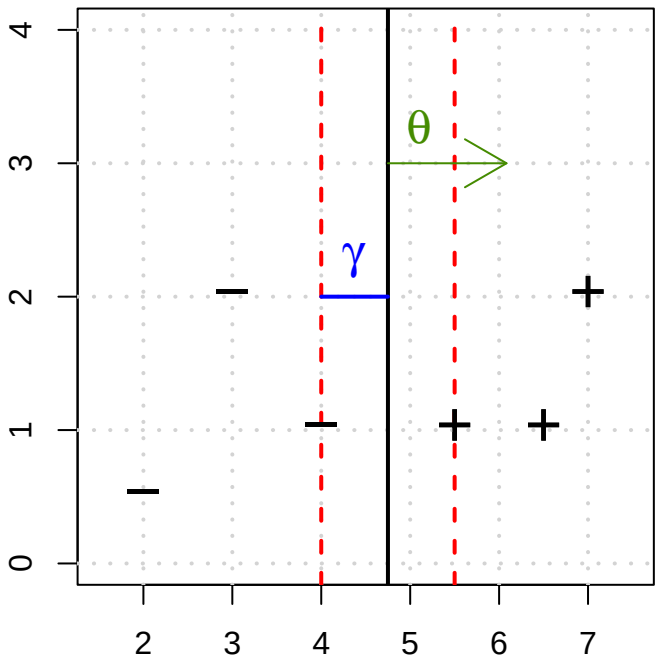
The safety margin γ and thus the norm of $\boldsymbol{\theta}$ remain the same. However, the direction of $\boldsymbol{\theta}$ changed because of the rotation. The support vectors are still the same as before, but rotated.

- We are now gonna shift all points by 2 to the right:



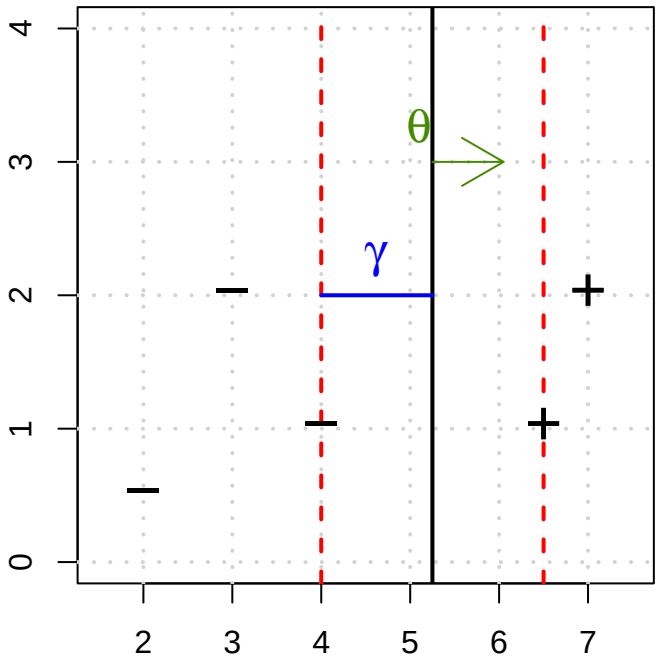
All magnitudes remain the same with the exception of θ_0, which is now -7. The support vectors are now shifted by 2 to the right.

- If one SV moves closer to the boundary $(6, 1) \rightarrow (5.5, 1)$:



The margin γ is now smaller, and thus $\|\boldsymbol{\theta}\|$ is bigger. The direction of $\boldsymbol{\theta}$ and θ_0 remain the same. The support vectors are now $(4,1)^T$ and $(5.5,1)^T$.

- If we remove a support vector such as $(6,1)^T$:



The margin γ increases, and thus $\|\boldsymbol{\theta}\|$ will decrease. The direction of $\boldsymbol{\theta}$ remains the same but θ_0 changes. The new Support vectors are $(4,1)^T$ and $(6.5,1)^T$.

Summarizing the four cases in one table:

γ	$\|\boldsymbol{\theta}\|$	$\boldsymbol{\theta}$	θ_0	SV
1	1	$\frac{1}{\sqrt{2}}(1,1)^T$	-5	Rotation $((4,1)^T,(6,1)^T)$
1	1	$(1,0)^T$	-7	$(6,1)^T,(8,1)^T$
0.75	4/3	$(4/3,0)^T$	-4.75	$(4,1)^T,(5.5,1)^T$
1.25	4/5	$(4/5,0)^T$	-5.25	$(4,1)^T,(6.5,1)^T$